

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-258881

(43)Date of publication of application : 11.09.2002

(51)Int.Cl.

G10L 11/02

G10L 15/04

G10L 15/20

G10L 21/02

(21)Application number : 2001-055324

(71)Applicant : FUJITSU LTD

(22)Date of filing : 28.02.2001

(72)Inventor : ENDOU KAORI
OTA TAKASHI
NAKAMA SHINJI

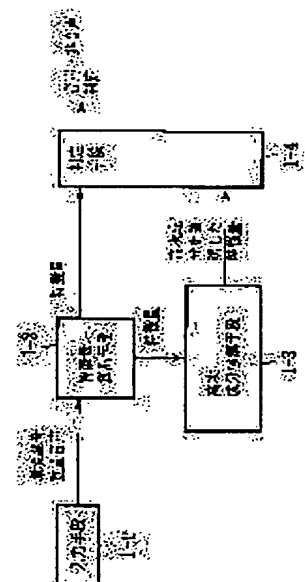
(54) DEVICE AND PROGRAM FOR DETECTING VOICE

(57)Abstract:

PROBLEM TO BE SOLVED: To decide voice/non-voice with high accuracy, even when environment noise is large around a voice-detecting device and a voice detection program which decide a voice/non-voice section from a voice signal with environment noise superimposed thereon.

SOLUTION: A voice characteristic quantity calculating means 1-2 calculates voice characteristic quantities, such as power of a voice signal, a zero crossing rate, peak frequency of a power spectrum and a pitch cycle, with respect to an input signal in each frame given time sequentially from an input means 1-1. A higher order component emphasizing means 1-3 calculates a characteristic quantity obtained by emphasizing the high pass component of a voice frequency characteristic quantity. A voice-deciding means 1-4 performs voice/non-voice decision, on the basis of the characteristic quantities. Furthermore, a means for calculating the existence/absence of the regularity of a voice resonant period is included to perform voice/non-voice decision, on the basis of the existence/ absence of the regularity.

本発明の音声検出装置の構成図



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

(19) 日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開 2002-258881

(P 2002-258881 A)

(43) 公開日 平成14年9月11日 (2002. 9. 11)

(51) Int. Cl. 7

識別記号

F I

テーマコード (参考)

G 1 0 L 11/02

G 1 0 L 3/00 5 1 3 A 5D015

15/04

3/02 3 0 1 D

15/20

21/02

審査請求 未請求 請求項の数 5

OL

(全 9 頁)

(21) 出願番号 特願2001-55324 (P2001-55324)

(22) 出願日 平成13年2月28日 (2001. 2. 28)

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番1号

(72) 発明者 遠藤 香緒里

神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

(72) 発明者 大田 恭士

神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

(74) 代理人 100105337

弁理士 眞鍋 潔 (外3名)

最終頁に続く

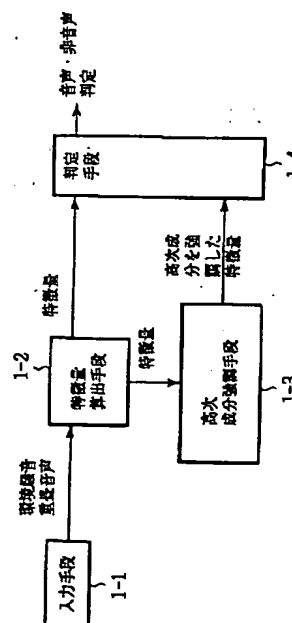
(54) 【発明の名称】 音声検出装置及び音声検出プログラム

(57) 【要約】

【課題】 環境騒音が重畳した音声信号から、音声・非音声区間を判定する音声検出装置及び音声検出プログラムに関し、環境騒音が大きい場合でも、精度良く音声・非音声を判定する。

【解決手段】 入力手段 1-1 から時系列的に順に与えられるフレーム毎の入力信号に対して、音声特徴量算出手段 1-2 により、音声信号の電力、零交差率、パワースペクトルのピーク周波数、ピッチ周期等の音声特徴量を算出する。高次成分強調手段 1-3 により、音声周波数の特徴量の高域成分を強調させた特徴量を算出する。音声判定手段 1-4 は、それらの特徴量を基に、音声・非音声判定を行う。更に、音声共振周期の規則性の有無を算出する手段を備え、該規則性の有無を基に音声・非音声判定を行う。

本発明の音声検出装置の第1の実施形態



【特許請求の範囲】

【請求項 1】 入力信号として環境騒音が重畳した音声信号が時系列順に与えられるフレーム毎に、当該フレームに音声が含まれているか否かを判定する音声検出装置であって、

電力、零交差率、パワースペクトルのピーク周波数、ピッチ周期のうち少なくとも一つを用いて音声の特徴量を算出する音声特徴量算出手段と、

パワースペクトルのピーク周波数の高次成分を強調して音声の特徴量を算出する高次成分強調特徴量算出手段と、

前記音声特徴量算出手段及び高次成分強調特徴量算出手段により算出された特徴量を基に、音声か非音声であるかを判定する音声判定手段と、

を備えたことを特徴とする音声検出装置。

【請求項 2】 入力信号として環境騒音が重畳した音声信号が時系列順に与えられるフレーム毎に、当該フレームに音声が含まれているか否かを判定する音声検出装置であって、

電力、零交差率、パワースペクトルのピーク周波数、ピッチ周期のうち少なくとも一つを用いて音声の特徴量を算出する音声特徴量算出手段と、

音声の共振周期の規則性を算出する共振周期の規則性算出手段と、

前記音声特徴量算出手段により算出した音声特徴量及び規則性算出手段により算出した音声の共振周期の規則性を基に、音声か非音声であるかを判定する音声判定手段と、

を備えたことを特徴とする音声検出装置。

【請求項 3】 入力信号として環境騒音が重畳した音声信号が時系列順に与えられるフレーム毎に、当該フレームに音声が含まれているか否かを判定する音声検出プログラムであって、

電力、零交差率、パワースペクトルのピーク周波数、ピッチ周期のうち少なくとも一つを用いて音声の特徴量を算出する音声特徴量算出手段と、

パワースペクトルのピーク周波数の高次成分を強調して音声の特徴量を算出する高次成分強調特徴量算出手段と、

前記音声特徴量算出手段及び高次成分強調特徴量算出手段により算出された特徴量を基に、音声か非音声であるかを判定する音声判定手段と、

を実行させる音声検出プログラム。

【請求項 4】 入力信号として環境騒音が重畳した音声信号が時系列順に与えられるフレーム毎に、当該フレームに音声が含まれているか否かを判定する音声検出プログラムであって、

電力、零交差率、パワースペクトルのピーク周波数、ピッチ周期のうち少なくとも一つを用いて音声の特徴量を算出する音声特徴量算出手段と、

音声の共振周期の規則性を算出する共振周期の規則性算出手段と、

前記音声特徴量算出手段により算出した音声特徴量及び規則性算出手段により算出した音声の共振周期の規則性を基に、音声か非音声であるかを判定する音声判定手段と、

を実行させる音声検出プログラム。

【請求項 5】 入力信号として環境騒音が重畳した音声信号が時系列順に与えられるフレーム毎に、当該フレームに音声が含まれているか否かを判定する音声検出プログラムであって、

電力、零交差率、パワースペクトルのピーク周波数、ピッチ周期のうち少なくとも一つを用いて音声の特徴量を算出する音声特徴量算出手段と、

パワースペクトルのピーク周波数の高次成分を強調して音声の特徴量を算出する高次成分強調特徴量算出手段と、

音声の共振周期の規則性を算出する共振周期の規則性算出手段と、

前記音声特徴量算出手段により算出した音声特徴量、高次成分強調特徴量算出手段により算出した特徴量、及び規則性算出手段により算出した音声の共振周期の規則性を基に、音声か非音声であるかを判定する音声判定手段と、

を実行させる音声検出プログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、音声検出装置及び音声検出プログラムに関する。一般に、音声信号を音声符号化器により圧縮して伝送する音声通信システム等では、通常の会話では約 6 割が発生の無い非音声区間であることに着目し、音声・非音声区間を判定することにより、非音声区間は伝送を停止して電池を長持ちさせたり、音声の代わりに他のデータを伝送するなどの通信制御を行うことができる。

【0002】音声通信システム等における音声検出機能は、このような通信制御を行なうために、音声・非音声区間の判定を行なう。しかし、騒音環境などの条件によっては音声・非音声区間の判定に誤りが起こり、それによって音切れが発生するなどして通話品質が著しく損なわれることがあり、音声・非音声区間を精度良く判定することが要求されている。

【0003】

【従来の技術】図 5 は従来の音声検出装置を示す。従来の音声検出装置は、環境騒音が重畳した音声信号が入力信号入力手段 5-1 に加えられ、入力手段 5-1 から時系列的に順に出力されるフレーム毎の入力信号に対して、音声特徴量算出手段 5-2 により、音声信号の電力、零交差率、パワースペクトルのピーク周波数、又はピッチ周期等の音声信号の特徴量を計算し、音声判定手

段5-3において、これらの音声特徴量を閾値と比較することによって、当該フレームに音声が含まれているかどうかを判定する。

【0004】

【発明が解決しようとする課題】しかしながら、従来の音声検出装置では環境騒音が大きい場合、雑音区間と音声区間との音声特徴量の差が小さくなり、音声・非音声区間を精度良く判定することが困難である。本発明は、環境騒音の影響を受け難くするように、音声特徴量に処理を加えることで、環境騒音が大きい場合でも、精度良く音声・非音声を判定することができる音声検出装置及び音声検出プログラムを提供することを目的とする。

【0005】

【課題を解決するための手段】本発明の音声検出装置は、(1)入力信号として環境騒音が重畳した音声信号が時系列順に与えられるフレーム毎に、当該フレームに音声が含まれているか否かを判定する音声検出装置であって、電力、零交差率、パワースペクトルのピーク周波数、ピッチ周期のうち少なくとも一つを用いて音声の特徴量を算出する音声特徴量算出手段と、パワースペクトルのピーク周波数の高次成分を強調して音声の特徴量を算出する高次成分強調特徴量算出手段と、前記音声特徴量算出手段及び高次成分強調特徴量算出手段により算出された特徴量を基に、音声か非音声であるかを判定する音声判定手段と、を備えたものである。

【0006】また、(2)電力、零交差率、パワースペクトルのピーク周波数、ピッチ周期のうち少なくとも一つを用いて音声の特徴量を算出する音声特徴量算出手段と、音声の共振周期の規則性を算出する共振周期の規則性算出手段と、前記音声特徴量算出手段により算出した音声特徴量及び規則性算出手段により算出した音声の共振周期の規則性を基に、音声か非音声であるかを判定する音声判定手段と、を備えたものである。

【0007】また、本発明の音声検出プログラムは、

(3)入力信号として環境騒音が重畳した音声信号が時系列順に与えられるフレーム毎に、当該フレームに音声が含まれているか否かを判定する音声検出プログラムであって、電力、零交差率、パワースペクトルのピーク周波数、ピッチ周期のうち少なくとも一つを用いて音声の特徴量を算出する音声特徴量算出手段と、パワースペクトルのピーク周波数の高次成分を強調して音声の特徴量を算出する高次成分強調特徴量算出手段と、前記音声特徴量算出手段及び高次成分強調特徴量算出手段により算出された特徴量を基に、音声か非音声であるかを判定する音声判定手段と、を実行させるものである。

【0008】また、(4)電力、零交差率、パワースペクトルのピーク周波数、ピッチ周期のうち少なくとも一つを用いて音声の特徴量を算出する音声特徴量算出手段と、音声の共振周期の規則性を算出する共振周期の規則性算出手段と、前記音声特徴量算出手段により算出した

音声特徴量及び規則性算出手段により算出した音声の共振周期の規則性を基に、音声か非音声であるかを判定する音声判定手段と、を実行させるものである。

【0009】また、(5)電力、零交差率、パワースペクトルのピーク周波数、ピッチ周期のうち少なくとも一つを用いて音声の特徴量を算出する音声特徴量算出手段と、パワースペクトルのピーク周波数の高次成分を強調して音声の特徴量を算出する高次成分強調特徴量算出手段と、音声の共振周期の規則性を算出する共振周期の規則性算出手段と、前記音声特徴量算出手段により算出した音声特徴量、高次成分強調特徴量算出手段により算出した特徴量、及び規則性算出手段により算出した音声の共振周期の規則性を基に、音声か非音声であるかを判定する音声判定手段と、を実行させるものである。

【0010】

【発明の実施の形態】図1は本発明の音声検出装置の第1の実施形態を示す。本発明の第1の実施形態は、入力手段1-1から時系列的に順に与えられるフレーム毎の入力信号に対して、音声特徴量算出手段1-2により、従来と同様に音声信号の特徴量を算出すると共に、高次成分強調手段1-3により、音声周波数の特徴量の高域成分を強調させた特徴量を算出する。音声判定手段1-4は、従来の特徴量に加えて高域成分を強調させた特徴量を基に、音声・非音声判定を行う。

【0011】従来の音声検出方式及び装置では、音声周波数の高次の成分に着目することなく、全般的なパワースペクトルのピーク周波数が音声と非音声とで異なることを利用して判定している。これに対し本発明では、音声は低域の周波数成分にエネルギーが集中するのに対し、非音声は一般に特定の周波数領域へのエネルギーの集中がないので、高次の成分ほど音声と非音声との特徴量の差が大きくなることを利用して、高次の成分の違いを強調した特徴量を算出する。この特徴量を用いて判定することで、騒音の影響が大きい場合でも、正確に音声・非音声を判定することができる。

【0012】図2は本発明の音声検出装置の第2の実施形態を示す。本発明の第2の実施形態は、入力手段2-1から時系列的に順に与えられるフレーム毎の入力信号に対して、音声特徴量算出手段2-2により、従来と同様に音声信号の特徴量を算出すると共に、共振周期の規則性算出手段2-3により、音声の共振周期の規則性を算出する。そして、音声判定手段2-4は、従来の特徴量に加えて音声の共振周期の規則性を基に、音声・非音声判定を行う。

【0013】従来の音声検出方式及び装置では、通常は最も自己相関が高くなる周期を共振周期として求め、該共振周期を用いて音声・非音声を判定する。しかし、騒音の種類によっては、音声と同じような共振周期を有する場合があります、このような場合には判別が困難となる。

【0014】母音などの音声は声道の共振によって生成

されるため、声道の基本周期の倍数にエネルギーが集中し、共振する複数の周期（共振周期）の間に規則性が現れるのに対し、非音声では特別な規則性がないことに着目し、本発明は、音声の共振周期の規則性を算出し、これを基に音声・非音声を判定する。

【0015】こうすることにより、最も自己相関が高くなる周期が音声と騒音で一致する場合でも、音声のみが複数の共振周期の間に規則性が現れるので、騒音と判定することができ、騒音の影響が大きい場合でも、正確に音声・非音声を判定することが可能になる。

【0016】更に、本発明の第1の実施形態として説明したパワースペクトルのピーク周波数の高次成分の違いを強調した特徴量と、第2の実施形態として説明した音声の共振周期の規則性とを組合わせて、音声・非音声の判定を行うことにより、騒音の影響が大きい場合でも、より正確に音声・非音声を判定することが可能となる。*

$$E_N = 0, Z_N = 0, LSF_N[i] = 0 \quad (1 \leq i \leq 10) \quad \dots (1)$$

【0019】次にステップ3-2により、入力音声を読み込む。次に、ステップ3-3により、現フレームのパラメータを以下の式(2)～(5)により算出する。ここで、現フレームの電力をE、零交差数をZ、共振周波数をF1と表す。

【数1】

$$E = 10 \log_{10} (1/N \sum_{i=1}^N x(i)^2) \quad \dots (2)$$

$$Z = (1/N \sum_{i=1}^N \text{sgn}(x(i)) \cdot x(i+1)) \quad \dots (3)$$

$$\text{corr}[k] = (\sum_{i=1}^N x(i) \cdot x(i-k)) / (\sum_{i=1}^N x(i-k) \cdot x(i-k)) \quad \dots (4)$$

$$F = \frac{s}{\text{MAX}(\text{corr}[k])} \quad \dots (5)$$

※

$$L1 = \sum_{i=1}^{10} (LSF[i] - LSF_N[i])^2 \quad \dots (7)$$

【0022】次にステップ3-5により、線スペクトル周波数の高次成分を強調した特徴量L2を算出する。高次成分を強調する手法の一つとして、最高次数の線スペクトル周波数だけを用い、以下の式(8)により高次成分を強調した特徴量L2を算出することができる。この算出処理は、本発明の第1の実施形態における高次成分強調特徴量算出手段で行われる算出処理に相当する。

$$L2 = LSF[10] - LSF_N[10] \quad \dots (8)$$

【0023】次にステップ3-6により、共振周期の規則性F2を算出する。この共振周期の規則性F2の算出処理は、本発明の第2の実施形態における共振周期の規

*【0017】図3は本発明の音声検出の処理フローを示す。同図はパワースペクトルのピーク周波数をPARCOR分析で得られる10次の線スペクトル周波数LSF[i] (1 ≤ i ≤ 10) とした場合の処理フローの例を示している。ここで、従来の音声検出により用いられる特徴量を、フレーム電力E、零交差数Z、10次の線スペクトル周波数から求めた特徴量L1、共振周波数F1とする。そして、パワースペクトルのピーク周波数の高次成分を強調した特徴量をL2、共振周期の規則性をF2とする。

【0018】先ず、ステップ3-1において、非音声区間のパラメータを以下の式(1)のとおり初期化する。ここで、非音声区間のフレーム電力をE_N、非音声区間の零交差数をZ_N、非音声区間の線スペクトル周波数をLSF_N[i]と表すものとする。

※【0020】ここで、x(i)は現フレームのi番目の入力音声信号振幅値であり、sgn(x(i))、x(i+1))は、x(i)とx(i+1)の符号が異なるとき1、同符号のとき0となる関数である。MAX(corr[k])は、k(1 ≤ k ≤ 143)サンプルずらして求めた自己相関のうち、最大値のkを表す。sはサンプリングレート(Hz)を表す。

【0021】次にステップ3-4により、現フレームの特徴量と非音声区間の特徴量の差分を以下の式(6)により算出する。ここで、フレーム電力の差分をΔE、零交差数の差分をΔZで表す。

$$\Delta E = E - E_N, \Delta Z = Z - Z_N \quad \dots (6)$$

また、L1を以下の式(7)により算出する。ここまでは、従来の音声特徴量算出処理と同様である。

【数2】

規則性算出手段で行われる算出処理に相当する。規則性F2の算出の詳細は、図4を参照して後述する。

【0024】次にステップ3-7により、非音声パラメータの更新を行うかどうかを判定するために、ΔE、ΔZ、L1、F1、L2、F2を、予め与えられた閾値TH_{EN}、TH_{ZN}、TH_{L1}、TH_{F1}、TH_{L2}、TH_{F2}とそれぞれ比較する。そして、
ΔE < TH_{EN} ∧ ΔZ < TH_{ZN} ∧ L1 < TH_{L1} ∧ F1 > TH_{F1} ∧ L2 < TH_{L2} ∧ F2 < TH_{F2}
の条件が成り立つ場合には更新を行うために、ステップ3-8へ処理を進め、上記条件が成り立たない場合に

は、更新を行うことなくステップ3-9へ処理を進める。

$$E_N = \alpha E + (1 - \alpha) E_N, Z_N = \beta Z + (1 - \beta) Z_N \quad \dots (9)$$

$$LSF[i]_N = \gamma LSF[i] + (1 - \gamma) LSF[i]_N \quad \dots (10)$$

α, β, γ は予め与えられた値で、 $0 \leq \alpha, \beta, \gamma \leq 1, 1 \leq i \leq 10$ である。

【0026】次にステップ3-9において、以下の式により音声・非音声の判定を行う。ここで、 $TH_E, TH_Z, TH_{L1}, TH_{F1}, TH_{L2}, TH_{F2}$ は予め与えられた閾値とする。

$$\Delta E > TH_E \quad U \Delta Z > TH_Z \quad U L1 > TH_{L1} \quad U F1 < TH_{F1} \quad U L2 > TH_{L2} \quad U F2 > TH_{F2}$$

の条件が成り立つ場合には、音声と判定してステップ3-10へ処理を進める。上記条件が成り立たない場合には非音声と判定してステップ3-11へ処理を進める。

【0027】ステップ3-10において音声と判定した※

*【0025】ステップ3-8において、非音声パラメータを以下の式(9)、(10)により更新する。

※旨の信号を出力する。ステップ3-11においては非音声と判定した旨の信号を出力する。ステップ3-12において次の入力信号が無いかどうかを判定し、無い場合には処理を終了する。次の入力信号が有る場合には、ステップ3-2に処理を進め、上記の処理フローを繰り返し実行する。

【0028】次に、図4を用いて本発明における共振周波数の規則性の算出処理を説明する。先ずステップ4-1において、偏自己相関値($corr[k]$)を以下の式(11)により計算する。 x は入力信号を示す。

【数3】

$$corr[k] = \frac{\sum_{j=1}^{80} x(j)x(j-k)}{\sum_{j=1}^{80} x(j-k)x(j-k)} \quad 1 \leq k \leq 143 \quad \dots (11)$$

【0029】ステップ4-2において、偏自己相関値のピーク値上位10個とそれに対応する k の値を取る。即ち、 $corr[k-1] < corr[k] > corr[k+1]$ となる偏自己相関値のうち、値が大きいものから上位10個を取り、これを $cor_max[i]$ とし、対応する k の値を $lag[i]$ $1 \leq i \leq 10$ とする。ステップ4-3において、 $cor_max[i]$ 及び $lag[i]$ を、 $lag[i]$ の昇順にソートする。ステップ4-4において、 $i=0, R=9, lnu$ 30
 $m=10$ とする。

【0030】次に、ピーク位置の間隔が所定値 $TH1$ 以★

$$cor_max[i] < cor_max[i+1] \quad \dots (13)$$

の式が成り立つか判定し、成り立つ場合ステップ4-8へ処理を進める。一方、

$$cor_max[i] \geq cor_max[i+1]$$

である場合、ステップ4-7へ処理を進める。

$$cor_max[j] = cor_max[j+1] \quad \dots (14)$$

$$lag[j] = lag[j+1] \quad i+1 \leq j \leq R-1 \quad \dots (15)$$

【0033】ステップ4-8において、

$$cor_max[i+1] / cor_max[i] < TH2 \quad \dots (16)$$

の式が成り立つか判定し、成り立つ場合ステップ4-1

0へ処理を進める。一方、 $cor_max[i+1] / cor_max[i] \geq TH2$ の場合、ステップ4-9へ処理を進める。

$$cor_max[j] = cor_max[j+1] \quad \dots (17)$$

$$lag[j] = lag[j+1] \quad i \leq j \leq R-1 \quad \dots (18)$$

【0035】ステップ4-10において、 cor_ma

$x[i+1], lag[i+1]$ を除去し、 cor_m *

$$cor_max[j] = cor_max[j+1] \quad \dots (19)$$

★下となるような近過ぎるもの同士を取り出し、その隣合ったもののうちピーク値の大きい方を採用し、所定の間隔以上のピーク値について規則性を調べるため、ステップ4-5において、

$$lag[i+1] - lag[i] < TH1 \quad \dots (12)$$

の式により、ピーク位置の間隔が所定値 $TH1$ 以下であるかどうか判定し、以下である場合ステップ4-6へ処理を進める。上記の式(12)が成り立たない、即ち所定の間隔 $TH1$ 以上の場合はステップ4-22へ処理を進める。

【0031】ステップ4-6において、

☆【0032】ステップ4-7において、 $cor_max[i+1], lag[i+1]$ を除去し、 cor_max, lag を以下の式(14)、(15)により更新し、ステップ4-11へ処理を進める。

◆【0034】ステップ4-9において、 $cor_max[i], lag[i]$ を除去し、 cor_max, lag を以下の式(17)、(18)により更新し、ステップ4-11へ処理を進める。

* ax, lag を以下の式(19)、(20)により更新し、ステップ4-11へ処理を進める。

$$\text{lag}[j] = \text{lag}[j+1] \quad i+1 \leq j \leq R-1 \quad \dots (20)$$

【0036】ステップ4-11において、 $R=R-1$ 、 $\text{lnum}=\text{lnum}-1$ とし、ステップ4-12へ処理を進める。また、前述のステップ4-22において、 $i=i+1$ とし、ステップ4-12へ処理を進める。ステップ4-12において $i < R$ かどうか判定し、 $i < R$ ならばステップ4-5へ処理を進める。 $i \geq R$ ならばステップ4-13へ処理を進める。

【0037】ステップ4-13において、 lag の数1

$$\text{TH3_1} < \text{lag}[1] / \text{lag}[0] < \text{TH3_2} \quad \dots (21)$$

が成り立つか判定し、成り立つ場合ステップ4-18へ処理を進める。 $\text{TH3_1} \geq \text{lag}[1] / \text{lag}[0]$

$[0]$ 、又は $\text{lag}[1] / \text{lag}[0] \geq \text{TH3_2}$

$$\text{TH4_1} < \text{lag}[1] / \text{lag}[0] < \text{TH4_2} \quad \dots (22)$$

$$\text{かつ、TH5_1} < \text{lag}[2] / \text{lag}[0] < \text{TH5_2} \quad \dots (23)$$

の条件が成り立つか判定し、成り立つ場合ステップ4-20へ処理を進める。一方、 $\text{TH4_1} \geq \text{lag}[1] / \text{lag}[0]$ 、又は $\text{lag}[1] / \text{lag}[0] \geq \text{TH4_2}$ 、又は $\text{TH5_1} \geq \text{lag}[2] / \text{lag}[0]$ 、又は $\text{lag}[2] / \text{lag}[0] \geq \text{TH5_2}$

である場合はステップ4-21へ処理を進める。

【0040】ステップ4-18では規則性有りを出力する。ステップ4-19では規則性無しを出力する。ステップ4-20では規則性有りを出力する。ステップ4-21では規則性無しを出力する。ここで、規則性有りの場合、共振周期の規則性F2の値として例えば1を、規則性無しの場合、共振周期の規則性F2の値として例えば0を出力する。

【0041】

【発明の効果】以上説明したように本発明によれば、従来の音声特徴量算出に加えて、環境騒音の影響を受け難いように、高次成分の違いを強調した音声特徴量の算出を行い、この音声特徴量を用いて音声・非音声を判定することにより、従来に比べて騒音の影響が大きい場合でも、より正確に音声・非音声を判定することができる。

【0042】また、従来の音声特徴量算出に加えて、音声共振周期の規則性の有無を算出し、この規則性を用いて音声・非音声を判定することにより、従来に比べて騒

num が2以上かどうか判定し、2以上ならばステップ4-15へ処理を進める。 lnum が1ならばステップ4-14へ処理を進める。ステップ4-14において、規則性なしを出力する。

【0038】ステップ4-15において、 lag の数が3以上かどうか判定し、3以上ならばステップ4-17へ処理を進める。 lag の数が2ならばステップ4-16へ処理を進める。ステップ4-16において、

である場合、ステップ4-19へ処理を進める。

【0039】ステップ4-17において、

音の影響が大きい場合でも、より正確に音声・非音声を判定することができる。

【0043】更に、従来の音声特徴量算出に加えて、環境騒音の影響を受け難いように高次成分の違いを強調した特徴量、及び音声共振周期の規則性の有無を算出し、これらを用いて音声・非音声を判定することにより、従来に比べて騒音の影響が大きい場合でも、より正確に音声・非音声を判定することができるようになる。

【図面の簡単な説明】

【図1】本発明の音声検出装置の第1の実施形態を示す図である。

【図2】本発明の音声検出装置の第2の実施形態を示す図である。

【図3】本発明の音声検出の処理フローを示す図である。

【図4】本発明における共振周波数の規則性の算出処理のフロー図である。

【図5】従来の音声検出装置を示す図である。

【符号の説明】

1-1 入力手段

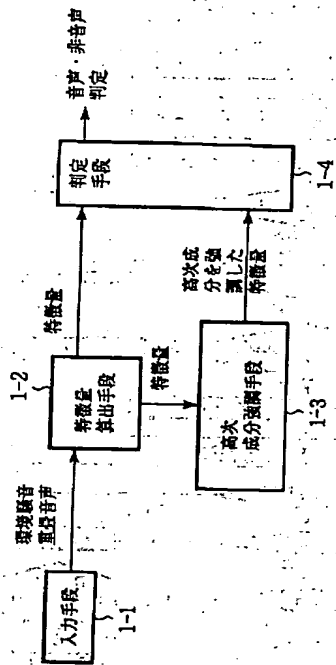
1-2 音声特徴量算出手段

1-3 高次成分強調手段

1-4 音声判定手段

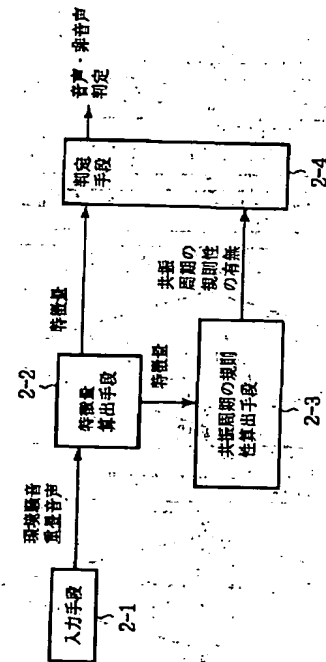
【図1】

本発明の音声検出装置の第1の実施形態

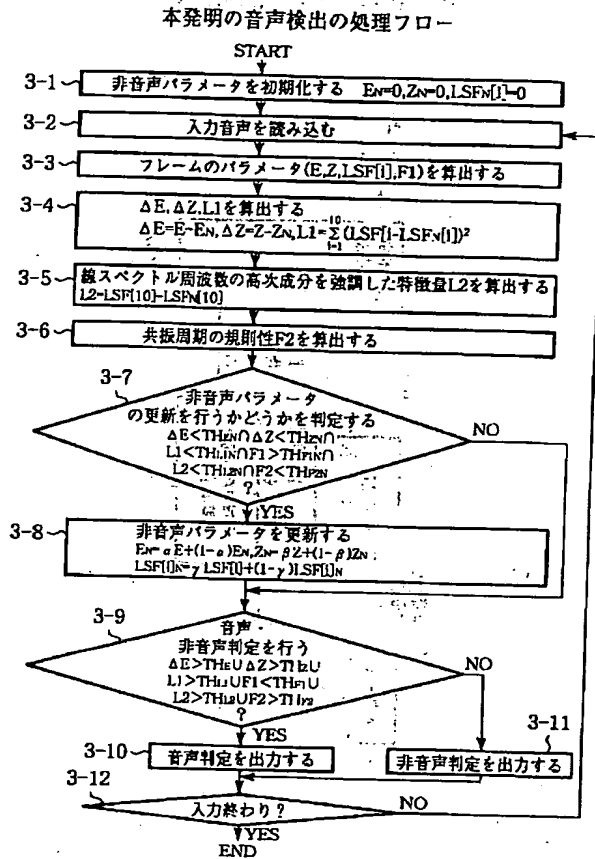


【図2】

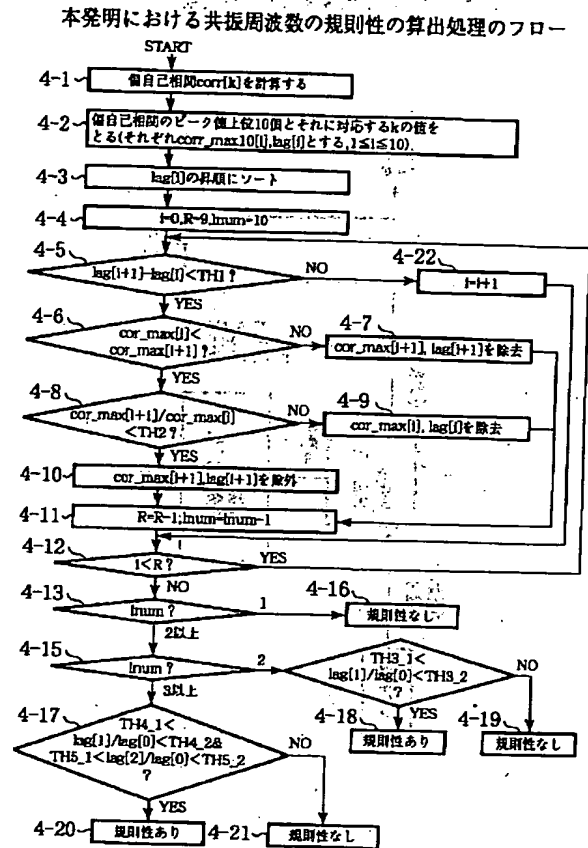
本発明の音声検出装置の第2の実施形態



【図3】

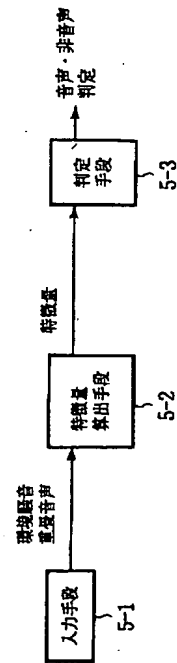


【図4】



【図5】

従来の音声検出装置



フロントページの続き

(72)発明者 中間 真司
 福岡県福岡市博多区博多駅前三丁目22番8
 号 富士通九州デジタル・テクノロジー株
 式会社内

Fターム(参考) 5D015 DD03 EE05